BMJ open
accessible medical research

Social inequality in breast, lung and colorectal cancers: a sibling approach

Grethe Søndergaard,[1] Laust Hvas Mortensen,,[2] Anne-Marie Nybo Andersen,[2] P K Andersen,[3] Susanne Oksbjerg Dalton,[4] Merete Osler[1]

To cite: Søndergaard G, Mortensen, LH, Andersen A-MN, *et al*. Social inequality in breast, lung and colorectal cancers: a sibling approach. *BMJ Open* 2013;**3**: e002114. doi:10.1136/bmjopen-2012-002114

▶ Prepublication history for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2012-002114).

Received 26 September 2012
Revised 25 January 2013
Accepted 28 January 2013

This final article is available for use under the terms of the Creative Commons Attribution Non-Commercial 2.0 Licence; see http://bmjopen.bmj.com

[1]Research Centre for Prevention and Health, Glostrup University Hospital, Copenhagen, Denmark
[2]Section for Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark
[3]Department of Biostatistics, University of Copenhagen, Copenhagen, Denmark
[4]Danish Cancer Society Research Center, Danish Cancer Society, Unit of Survivorship, Copenhagen, Denmark

Correspondence to:
Dr Merete Osler;
merete.osler@regionh.dk

ABSTRACT

Objective: To examine whether family factors shared by siblings explained the association between education and risk of lung, colorectal and breast cancer.

Design: We used conventional cohort and intersibling Cox regression analyses to analyse the association between education and risk of cancer.

Setting: Denmark.

Participants: We retrieved register data from Statistics Denmark on individuals born in Denmark 1950–1979 with at least one full sibling. The cohorts included between 391 931 and 1 381 369 individuals followed from age 28 for incident lung, colorectal and breast cancer until the end of 2009.

Results: In the cohort analysis, low education was associated with an increased risk of colorectal cancer before age 45 and lung cancer, and with a decreased risk of colorectal cancer after age 45 and breast cancer. When compared with the cohort analyses, the intersibling associations were stronger for colorectal cancer after age 45 and weaker for lung cancer. Serious health conditions in childhood/young adulthood did not explain the associations.

Conclusions: Family factors shared by siblings confounded some of the association between education and colorectal cancer after age 45 and lung cancer, but not the associations found for colorectal cancer before age 45 or breast cancer.

ARTICLE SUMMARY

Article focus

- Whether the family environment in childhood and genes shared by siblings explain the association between educational status and lung, colorectal and breast cancer, respectively.

Key messages

- Low education was associated with an increased risk of colorectal cancer before age 45 and lung cancer and with a decreased risk of colorectal cancer after age 45 and breast cancer.
- Serious health conditions in childhood/young adulthood did not explain the associations.
- Intersibling comparisons indicated that family factors shared by siblings confound some of the association between education and colorectal cancer after age 45 and lung cancer, but not the associations for colorectal cancer before age 45 or breast cancer.

Strengths and limitations of this study

- The study is based on data from a large unselected population, using relatively valid information from the Danish registers.
- Parental links for persons born before 1968 may not be precise or obtainable in all cases.
- The large differences in educational status found between some of the siblings may be caused by factors not shared, and it is likely that residual confounding from unobserved unshared causes occur.

INTRODUCTION

A high risk of lung cancer has consistently been associated with a low educational status,[1 2] while a positive association has been found for breast cancer.[1–5] Studies show, however, mixed results for colorectal cancer.[6] A relationship between educational status and these cancers might be explained by social differences in risk factors. It has been found that the social gradient in lung cancer is partly because of a higher smoking prevalence among the lowest educational groups,[7] while postponement and reduction of child births,[1] a higher alcohol consumption and use of hormone replacement therapy[8] among higher educated may explain the positive association for breast cancer. The contradictory findings for colorectal cancer indicate that it is explained by factors that vary across time and populations. The social inequalities may also be explained by family characteristics such as genetics and family environment in childhood, affecting both achieved educational status and health in adulthood.[9–11] A study has found that childhood socioeconomic status was not associated with breast or lung cancer when socioeconomic status in adulthood was adjusted for but that this was the case for colorectal cancer.[12] However, important factors linked to family characteristics such as genetic factors may not be included in such studies which mean that the estimated association will still not reflect the causal effect.[13 14]

Another approach is to examine the association between educational status and health within siblings discordant on educational status. Siblings are matched on the family environment and genetic factors that they share, making it possible to adjust for these per design. Insofar confounding occurs the association will be attenuated compared with the association between non-siblings. A twin design has been used to examine the association between socioeconomic status and breast cancer[3] but the study was limited owing to a small sample size and high concordance in exposure.

The aim of this study was to examine if family environment in childhood and genes shared by siblings explained the association between educational status and lung, colorectal and breast cancer, respectively, among full siblings born in Denmark between 1950 and 1979.

MATERIAL AND METHODS

Study population

The current study was based on register linkage of data from the Central Population Register, the Integrated Database for Labour Market Research, the Cancer Registry, the medical birth- and death statistics, and the Central Danish Psychiatric Register, using the unique identification number assigned to all Danish citizens. The study population included three cohorts, each of which consisted of individuals born in Denmark between 1950 and 1979 who had at least one full sibling born during the same period and who had not died or emigrated before age 28. Cohort 1 included individuals not diagnosed with lung cancer before age 28 (N=1 381 369), cohort 2 individuals not diagnosed with colorectal cancer before age 28 (N=1 381 266) and cohort 3 women who were not diagnosed with breast cancer before age 28 and who had at least one female sibling born during 1950–1979 (N=391 931). Full siblings were defined as persons registered with the same father and mother, thus cohorts 1 and 2 consisted of siblings from approximately 567 000 families, and cohort 3 of siblings from about 178 000 families. About 1.5% was excluded in the analysis because of missing information on educational status and 2–3% emigrated during follow-up.

Exposure and outcome

Education was categorised into five ordinal groups based on type and level of education and ranged from 'primary school' (no education beyond mandatory schooling), 'high school' (A level, International Baccalaureate), 'vocational education' (eg carpenters, salesperson or hairdressers with a technical school education), 'short and middle-length higher education' (eg BA, primary school teachers and nurses), and 'advanced higher education' (eg MSc, MD or PhD). The information on education was recorded on the 1 January in the year the person turned 28 since we assumed that most individuals had reached or started their highest level of education at that age. Information about date of diagnosis of lung, colorectal and breast cancer was retrieved from the Cancer Registry using International Classification of Diseases (ICD-8) (1981–1994) and ICD-10 (1994–). Codes 170 and C50 were used for breast cancer, 162.0,1,8 and C33–34 for lung cancer and 153–4 and C18–21 for colorectal cancer.

Covariates

Sex was included as potential confounder in the analyses of lung and colorectal cancer. Maternal breast cancer (ever/never diagnosed) was included as a potential confounder in the analysis of breast cancer to examine whether family history of breast cancer accounted for the association. To examine if low health status during childhood and young adulthood caused a low educational status at age 28 and affected cancer risk later in life, we included psychiatric hospitalisation in young adulthood (ever/never) and disability pension at age 28 as an indicator of serious health conditions. All information was retrieved from the Medical Register of Birth and Deaths, the Central Danish Psychiatric Register and the Integrated Database for Labour market Research. Because of limitations in register coverage, information about psychiatric hospitalisation was only retrieved for the 19–27 years old.

Data analysis

The risk of cancer was estimated through survival analysis using Cox regression models with age as the underlying time variable. Person-years of follow-up were calculated from 1 January in the year the person turned 28 until the age of the cancer diagnosis, emigration or the end of follow-up (31 December 2009), whichever came first. Initially, conventional cohort analyses were made, treating all siblings as individuals while correcting the standard errors using the robust sandwich variance estimator to take the interdependence between siblings into account. The results from this analysis were compared with the results of a sibling analysis in which we included a sibling stratum variable, taking the same value for individuals being full siblings. This means that a separate baseline hazard is estimated for every sibling group.[15] Thus, the HR were estimated relative to the family-level baseline hazard, and are as such similar to conditional logistic regression. Hence, it was possible to compare siblings, keeping fixed the shared familial factors that acted multiplicatively on the hazard function within siblings, thereby controlling for these per design. The difference between the cohort and intersiblings estimates was used as a crude measure of the degree of confounding. An attenuation of associations in the intersibling analysis will reflect confounding because of shared familial factors.[16] Separate estimates were obtained for men and women, whereby only same sex siblings contributed to the estimations. We found no statistically significant interactions between education and birth year or between education and sex for the three outcomes. The proportional hazards assumption was violated in the analysis for colorectal cancer, hence we analysed the association between educational status and colorectal before

and after age 45 separately. Age 45 was chosen as the cut-off point because of potential differences in aetiology before and after this age and to ensure that even enough events occurred in each model. The proportional hazard assumption was not violated in the analysis of breast cancer, but we analysed this association separately for breast cancers before and after age 45 to provide separate estimates before and after menopause. To examine if a linear trend existed over the educational categories, likelihood ratio tests were carried out, comparing Cox regression models treating educational status as a categorical and quantitative variable, respectively.

RESULTS

Descriptive statistics

A total of 1456 lung, 1789 colorectal and 3526 breast cancers were recorded during follow-up and approximately 46% of lung, 53% of colorectal (of which 59% were in colon), and 63% of breast cancers appeared before age 45. The high number of cancers before age 45 was because of the fact that the study populations were 28–59 years at the end of follow-up. Table 1 presents the incidence rates and characteristics of cohorts 1 and 3 according to education while cohort 2, which almost alike cohort 1 is not presented. The incidence

Table 1 Incidence rates and characteristics of persons born in Denmark between 1950 and 1979 with at least one full sibling born during the same time period according to educational status in cohorts 1 and 3

	Total (N)*	Primary education (%)	High school (%)	Vocational education (%)	Short and middle-length higher education (%)	Advanced higher education (%)
COHORT 1*	*1 361 117*	*22.6*	*6.6*	*38.7*	*22.1*	*9.9*
Number of siblings						
1–2	1 167 946	20.7	6.9	39.1	22.8	10.5
3–4	174 985	32.7	5.2	37.3	18.1	6.7
5–6	15 270	47.5	3.8	32.4	12.7	3.6
>6	2916	54.6	2.8	30.3	10.2	2.1
Birth year						
1950–1959	267 562	28.5	5.2	36.0	21.4	8.9
1960–1969	622 525	24.4	7.2	40.8	19.1	8.5
1970–1979	471 030	17.0	6.6	37.6	26.4	12.4
Conditions in early life						
Disability pension	13 674	85.3	4.7	7.8	1.5	0.7
Psychiatric hospitalisation	21 380	50.2	8.7	25.8	11.5	3.9
Incidence rates						
Number of lung cancer diagnoses (N)	1415	592	65	507	190	61
Person-years (N)	21 059 716	5 186 286	1 345 361	8 188 156	4 430 641	1 909 273
Rate pr. 1000	0.1	0.1	0.1	0.1	0.0	0.0
COHORT 3†	*386 839*	*23.0*	*7.4*	*34.0*	*26.6*	*9.0*
Number of siblings						
1–2	369 728	22.4	7.4	34.1	26.9	9.1
3–4	16 382	35.3	6.4	31.8	21.0	5.5
5–6	630	51.9	4.0	26.7	14.9	2.5
>6	99	48.5	2.0	38.4	10.1	1.0
Birth year						
1950–1959	66 380	31.2	5.3	27.2	29.2	7.1
1960–1969	188 078	25.4	8.3	36.7	22.2	7.5
1970–1979	132 381	15.6	7.1	33.7	31.6	12.1
Conditions in early life						
Disability pension	3851	86.1	4.9	6.3	1.9	0.8
Psychiatric hospitalisation	5432	45.8	10.1	23.0	17.0	4.1
Incidence rates						
Number of breast cancer diagnoses (N)	3471	905	239	1112	941	274
Person-years (N)	5 932 111	1 526 086	423 680	1 992 699	1 522 335	467 311
Rate pr. 1000	0.6	0.6	0.6	0.6	0.6	0.6

*Individuals not diagnosed with lung cancer before age 28. In total 20 252 individuals are excluded because of missing information on education.
†Individuals not diagnosed with breast cancer before age 28. In total 5092 individuals are excluded because of missing information on education.

rates for breast and colorectal cancer did not vary substantially across the educational groups, while the rates ranged from 0.03 to 0.11 for lung cancer, being the highest for the lowest educated. Approximately 56% of the siblings were discordant on educational status in each cohort.

Lung cancer

In both the cohort and the intersibling analyses a lower educational status was associated with a higher risk of lung cancer (Table 2). In the cohort analysis, individuals with primary school education had an increased risk of lung cancer (HR=1.63 (95% CI 1.45 to 1.84)) compared with those with vocational education. The association attenuated in the intersibling analysis (HR=1.24 (95% CI 1.00 to 1.52)) but a significant trend persisted (HR=0.89 (95% CI 0.82 to 0.96)).

Colorectal cancer

In the cohort analysis a lower educational status was associated with an increased risk of colorectal cancer before age 45, whereas the opposite was found for colorectal cancer after age 45 (Table 3). The association for colorectal cancer after age 45 strengthened in the intersibling analyses, resulting in a significant trend estimate. Compared with those with a vocational education, the risk of colorectal cancer after 45 was less pronounced among individuals with advanced higher education in the cohort analysis (HR=1.11 (95% CI 0.87 to 1.41) than in the intersibling analysis (HR=1.76 (95% CI 1.05 to 2.95; table 3).

Breast cancer

The risk of breast cancer increased slightly with increasing educational status in the unadjusted cohort analysis (table 4). Adjusting the cohort analysis for maternal breast cancer did not change the associations nor did the associations change considerably in the intersibling analysis. The trend estimates turned insignificant in the intersibling analyses.

Adjusting for disability pension at age 28 and hospitalisation for a psychiatric illness did not influence the estimates in any of the analyses.

In general, the age difference between siblings (more/less than 5 years) and living together with the same parent(s) at age 15 years (yes/no) did not affect the associations. Only individuals with one sibling were, however, included in these subanalyses and because of register-coverage and number of cases, the last mentioned analysis was only performed on breast cancer. When stratifying on family size (two siblings versus more than two siblings) no differences in associations were found for lung and colorectal cancer but minor differences were found for breast cancer. Individuals excluded because of missing information on education had a slightly increased risk of colorectal cancer in the cohort analysis and of lung cancer, whereas they had a reduced risk of breast cancer after age 45.

DISCUSSION

In this cohort study of all siblings born in Denmark between 1950 and 1979 a lower educational status was associated with a higher risk of both colorectal cancer before age 45 and especially lung cancer whereas a higher educational status was associated with an increased risk of colorectal cancer after age 45 and breast cancer. The association attenuated for lung cancer in the intersibling analysis, indicating that factors shared by siblings explained a part of the associations. Contrary to this, the association between educational status and colorectal cancer after age 45 increased in the intersibling analysis whereas no remarkable changes appeared for colorectal cancer before age 45 and for breast cancer. Indicators of serious health conditions did not seem to explain the associations. The non-genetic family factors that are matched out in the intersibling analyses will potentially include known major risk factors such as smoking, diet, obesity and physical activity to the extent that it is shared between the siblings, possibly because the siblings' propensity to become exposed to these risk factors is shared. Because the genetic and non-genetic factors are unmeasured it is impossible

Table 2 Unadjusted HRs with 95% confidence limits for lung cancer according to educational status among full siblings born in Denmark between 1950 and 1979, showing the results from the cohort and intersibling analyses

	Cohort analysis*		Sibling analysis*	
Age at diagnosis/education	**HR**	**95% CI**	**HR**	**95% CI**
All age groups				
Primary school	1.64	1.45 to 1.84	1.24	1.00 to 1.54
High school	0.82	0.63 to 1.06	1.02	0.67 to 1.54
Vocational education	1		1	
Short and middle-length higher education	0.66	0.55 to 0.77	0.86	0.65 to 1.15
Advanced higher education	0.52	0.40 to 0.68	0.69	0.44 to 1.07
Trend estimates†	0.76	0.73 to 0.79	0.89	0.82 to 0.96

*Adjusted for sex.
†For every step up the educational ladder.

Table 3 Unadjusted HRs with 95% confidence limits for colorectal cancer according to educational status among full siblings born in Denmark between 1950 and 1979, showing the results from the cohort and intersibling analyses

	Cohort analysis*		Sibling analysis*	
Age at diagnosis/education	**HR**	**95% CI**	**HR**	**95% CI**
Colorectal cancer <age 45				
Primary school	1.16	0.98 to 1.36	1.28	0.99 to 1.64
High school	1.13	0.86 to 1.47	1.18	0.78 to 1.79
Vocational education	1		1	
Short and middle-length higher education	1.02	0.85 to 1.21	0.93	0.69 to 1.24
Advanced higher education	0.89	0.68 to 1.15	0.87	0.57 to 1.32
Trend estimates†	0.94	0.89 to 0.99	0.90	0.82 to 0.98
Colorectal cancer ≥ age 45				
Primary school	0.97	0.82 to 1.15	0.72	0.51 to 1.00
High school	0.81	0.57 to 1.14	0.82	0.47 to 1.43
Vocational education	1		1	
Short and middle-length higher education	0.93	0.77 to 1.13	1.02	0.71 to 1.47
Advanced higher education	1.11	0.87 to 1.41	1.76	1.05 to 2.95
Trend estimates†	1.02	0.96 to 1.07	1.19	1.06 to 1.33

*Adjusted for sex.
†For every step up the educational ladder.

from this study to assess the importance of any specific risk factor or the relative importance of the genetic and non-genetic factors.

Lung cancer

The negative association between educational status and lung cancer may be caused by higher smoking prevalence among the lowest educated.[7] In our study, the association attenuated in the intersibling analysis, indicates that family factors confounded the association. Opposite this, a study by de Kok *et al*[12] found that childhood socioeconomic status was not associated with lung cancer after adjusting for educational status in adulthood. However, childhood socioeconomic status might

Table 4 Unadjusted and adjusted HRs with 95% confidence limits for breast cancer according to educational status among full siblings born in Denmark between 1950 and 1979, showing the results from the cohort and intersibling analyses

	Cohort analysis		Cohort analysis*		Sibling analysis	
Age at diagnosis/education	**HR**	**95% CI**	**HR**	**95% CI**	**HR**	**95% CI**
All age groups						
Primary school	0.92	0.84 to 1.00	0.92	0.84 to 1.01	0.96	0.83 to 1.12
High school	1.04	0.90 to 1.19	1.02	0.89 to 1.18	0.99	0.80 to 1.26
Vocational education	1		1		1	
Short and middle-length higher education	1.04	0.95 to 1.13	1.03	0.94 to 1.12	1.03	0.89 to 1.20
Advanced higher education	1.10	0.97 to 1.26	1.08	0.95 to 1.24	1.04	0.82 to 1.33
Test for trend†	1.04	1.01 to 1.07	1.04	1.01 to 1.06	1.02	0.97 to 1.07
Cancer <age 45						
Primary school	0.91	0.82 to 1.02	0.91	0.82 to 1.02	0.98	0.82 to 1.18
High school	1.01	0.85 to 1.20	1.00	0.84 to 1.18	1.00	0.78 to 1.31
Vocational education	1		1		1	
Short and middle-length higher education	1.07	0.96 to 1.19	1.06	0.95 to 1.18	1.05	0.88 to 1.26
Advanced higher education	1.07	0.90 to 1.26	1.05	0.89 to 1.24	1.08	0.81 to 1.45
Test for trend†	1.05	1.01 to 1.08	1.04	1.01 to 1.08	1.02	0.96 to 1.08
Cancer ≥ age 45						
Primary school	0.93	0.81 to 1.07	0.93	0.81 to 1.08	0.91	0.69 to 1.21
High school	1.1	0.86 to 1.40	1.08	0.85 to 1.38	0.99	0.60 to 1.62
Vocational education	1		1		1	
Short and middle-length						
Higher education	1	0.86 to 1.15	0.99	0.86 to 1.14	0.97	0.72 to 1.30
Advanced higher education	1.16	0.93 to 1.45	1.15	0.92 to 1.43	0.95	0.61 to 1.49
Trend estimates†	1.03	0.99 to 1.08	1.03	0.99 to 1.07	1.02	0.93 to 1.12

*Adjusted for maternal breast cancer.
†For every step up the educational ladder.

not reflect the entire width of genetic and shared family environmental factors such as smoking, which the present sibling study accounts for.

Colorectal cancer

Similar to our result, a study also showed that the socioeconomic gradient in colorectal cancer mortality and stage at diagnosis varied across age groups and was strongest for younger individuals aged 40–59[17] indicating difference in aetiologies. A study found that higher social position was related to early stage at diagnosis in rectal but not in colon cancer and a more rapid response to symptoms of cancer among higher educated[18] may explain the differing social gradients in our study population. The low number of cases of colorectal cancer meant that we could not perform separate analyses for colon and rectal cancers. Only the siblings that are different with regard to education are informative for intersibling analyses, whereas the cohort associations are based on everyone who met the inclusion criteria. The stronger associations for colorectal cancer after age 45 in the intersibling analysis may reflect that differences in non-shared risk factors are more pronounced between siblings that end up with different levels of educational attainment than between unrelated individuals. Without information on such non-shared risk factors it is difficult to know exactly what causes the gradient in the intersiblings analyses to be stronger than the cohort gradient.

Breast cancer

As found in previous studies[1–5] educational status was associated with breast cancer. A study found that the association was slightly stronger for postmenopausal breast cancers, which was not found in our study. Nor did the associations change remarkably in the intersibling analysis, suggesting no confounding by family environment and genetic factors. In accordance with this, the previously mentioned study by de Kok *et al*[12] found that childhood socioeconomic status was not associated with breast cancer after adjusting for educational status in adulthood. However, a Danish twin study found that genetic factors probably confounded the association between educational status and breast cancer.[3]

Strengths and limitations

The present study is based on data from a large unselected population, using relatively valid information from the Danish registers. However, parental links for persons born before 1968 may not be precise or obtainable in all cases.[19] Sensitivity analyses showed that individuals excluded because they were lacking information on education had a slightly increased risk of colorectal cancer in the cohort analysis and of lung cancer but a decreased risk of breast cancer after age 45. We believe that such a selection would tend to underestimate the association between educational status and lung and colorectal cancers, respectively, whereas the association between educational status and breast cancers may be overestimated.

The large differences in educational status found between some of the siblings may be caused by factors not shared. Mental and physical limitations did, however, not explain the associations between educational status and the three cancer types but it is likely that residual confounding from unobserved unshared causes occur.

Age difference and whether siblings lived together during childhood are factors that determine to which extent siblings have been exposed to the same family environment. Only minor differences were found in a subanalysis examining the association between educational status and breast cancer on siblings who lived together with the same parent(s) at age 15 compared with the association based on all siblings. Age-difference did not seem to change the associations.

In conclusion, this study found that shared family environment and genetic factors explained some of the association for lung cancer but not for breast or for colorectal cancers before age 45. Contrary to this, the association between educational status and colorectal cancer after age 45 increased in the intersibling analysis.

Contributors All authors have contributed to the study. MO formulated the original study idea and provided funding. GS further developed the study idea and design of the study in collaboration with LHM, MO and A-MNA. GS carried out the data analyses and took the responsibility of writing the first draft of the paper. All authors participated in interpretation and discussion of the findings and writing of the final version of the paper.

Funding This work was supported by the Danish Cancer Society (Grant SU08001). LHM was supported by an individual postdoc grant from the Danish Council for Independent Research | Medical Sciences (Grant 09–066343).

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Vidarsdottir H, Gunnarsdottir HK, Olafsdottir EJ, *et al*. Cancer risk by education in Iceland; a census-based cohort study. *Acta Oncol* 2008;47:385–90.
2. Sidorchuk A, Agardh EE, Aremu O, *et al*. Socioeconomic differences in lung cancer incidence: a systematic review and meta-analysis. *Cancer Causes Control* 2009;20:459–71.
3. Madsen M, Andersen PK, Gerster M, *et al*. Does the association of education with breast cancer replicate within twin pairs? A register-based study on Danish female twins. *Br J Cancer* 2011;104:520–3.
4. Hussain SK, Altieri A, Sundquist J, *et al*. Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. *Int J Cancer* 2008;122:165–9.
5. Carlsen K, Hoybye MT, Dalton SO, *et al*. Social inequality and incidence of and survival from breast cancer in a population-based study in Denmark, 1994–2003. *Eur J Cancer* 2008;44:1996–2002.
6. Aarts MJ, Lemmens VE, Louwman MW, *et al*. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. *Eur J Cancer* 2010;46:2681–95.
7. Louwman WJ, van Lenthe FJ, Coebergh JW, *et al*. Behaviour partly explains educational differences in cancer incidence in the south-eastern Netherlands: the longitudinal GLOBE study. *Eur J Cancer Prev* 2004;13:119–25.
8. Larsen SB, Olsen A, Lynch J, *et al*. Socioeconomic position and lifestyle in relation to breast cancer incidence among

postmenopausal women: a prospective cohort study, Denmark, 1993–2006. *Cancer Epidemiol* 2011;35:438–41.
9. Lichtenstein P, Harris JR, Pedersen NL, *et al*. Socioeconomic status and physical health, how are they related? An empirical study based on twins reared apart and twins reared together. *Soc Sci Med* 1993;36:441–50.
10. Smith GD, Blane D, Bartley M. Explanations for socio-economic differences in mortality. *Eur J Public Health* 1994;4:131–44.
11. Marmot M, Ryff CD, Bumpass L, *et al*. Social inequalities in health: next questions and converging evidence. *Soc Sci Med* 1997;44:901–10.
12. de Kok IM, van Lenthe FJ, Avendano M, *et al*. Childhood social class and cancer incidence: results of the globe study. *Soc Sci Med* 2008;66:1131–9.
13. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. *J Health Econ* 2005;24:365–89.
14. Haas SA. Health selection and the process of social stratification: the effect of childhood health on socioeconomic attainment. *J Health Soc Behav* 2006;47:339–54.
15. Holt JD, Prentice RL. Survival Analysis in Twin Studies and Matched Pair Experiments. *Biometrika* 1974;61:17–30.
16. Janes H, Dominici F, Zeger S. On quantifying the magnitude of confounding. *Biostatistics* 2010;11:572–82.
17. Marcella S, Miller JE. Racial differences in colorectal cancer mortality. The importance of stage and socioeconomic status. *J Clin Epidemiol* 2001;54:359–66.
18. Frederiksen BL, Osler M, Harling H, *et al*. Social inequalities in stage at diagnosis of rectal but not in colonic cancer: a nationwide study. *Br J Cancer* 2008;98:668–73.
19. Pedersen CB, Gotzsche H, Moller JO, *et al*. The Danish Civil Registration System. A cohort of eight million persons. *Dan Med Bull* 2006;53:441–9.